Statistics and Probability Letters 139 (2018) 119–128

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The distance between a naive cumulative estimator and its least concave majorant

Hendrik P. Lopuhaä, Eni Musta *

DIAM, Faculty EEMCS, Delft University of Technology, van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands

ARTICLE INFO

Article history:
Received 16 June 2017
Received in revised form 23 November 2017
Accepted 2 April 2018
Available online 11 April 2018

MSC:
60F05
62E20

Keywords:
Least concave majorant
Grenander-type estimator
Limit distribution
Central limit theorem for L_p-distance
Brownian motion with parabolic drift

ABSTRACT

We consider the process $\widehat{\Lambda}_n - \Lambda_n$, where Λ_n is a cadlag step estimator for the primitive Λ of a nonincreasing function λ on $[0, 1]$, and $\widehat{\Lambda}_n$ is the least concave majorant of Λ_n. We extend the results in Kulikov and Lopuhaä (2006, 2008) to the general setting considered in Durot (2007). Under this setting we prove that a suitably scaled version of $\widehat{\Lambda}_n - \Lambda_n$ converges in distribution to the corresponding process for two-sided Brownian motion with parabolic drift and we establish a central limit theorem for the L_p-distance between $\widehat{\Lambda}_n$ and Λ_n.

1. Introduction

Grenander-type estimators are well known methods for estimation of monotone curves. In case of estimating nonincreasing curves, they are constructed by starting with a naive estimator for the primitive of the curve of interest and then take the left-derivative of the least concave majorant (LCM) of the naive estimator. The first example can be found in Grenander (1956) in the context of estimating a nonincreasing density f on $[0, \infty)$ on the basis of an i.i.d. sample from f. The empirical distribution function F_n of the sample is taken as a naive estimator for the cumulative distribution function corresponding to f and the Grenander estimator is found by taking the left-derivative $\widehat{f}_n$ of the least concave majorant $\widehat{F}_n$. Similar estimators have been developed in other statistical models, e.g., regression (see Brunk, 1958), random censoring (see Huang and Wellner, 1995), or the Cox model (see Lopuhaä and Nane, 2013). Durot (2007) considers Grenander-type estimators in a general setup that incorporates several statistical models. A large part of the literature is devoted to investigating properties of Grenander-type estimators for monotone curves, and somewhat less attention is paid to properties of the difference between the corresponding naive estimator for the primitive of the curve and its LCM.

Kiefer and Wolfowitz (1976) show that $\sup_t |\widehat{F}_n - F_n| = O_p((n^{-1}\log n)^{2/3})$. Although the first motivation for this type of result has been asymptotic optimality of shape constrained estimators, it has several important statistical applications. The Kiefer–Wolfowitz result was a key argument in Sen et al. (2010) to prove that the m out of n bootstrap from $\widehat{F}_n$ works. Mammen (1991) suggested to use the result to make an asymptotic comparison between a smoothed Grenander-type estimator and an isotonized kernel estimator in the regression context. See also Wang and Woodroofe (2007) for a similar application of their Kiefer–Wolfowitz comparison theorem. An extension to a more general setting was established

* Corresponding author.
E-mail address: e.musta@tudelft.nl (E. Musta).

https://doi.org/10.1016/j.spl.2018.04.001
0167-7152/

in Durot and Lopuhaä (2014), which has direct applications in Durot et al. (2013) to prove that a smoothed bootstrap from a Grenander-type estimator works for k-sample tests, and in Groeneboom and Jongbloed (2013) and Lopuhaä, and Musta (2017) to extract the pointwise limit behavior of smoothed Grenander-type estimators for a monotone hazard from that of ordinary kernel estimators. To approximate the L_p-error of smoothed Grenander-type estimators by that of ordinary kernel estimators, such as in Csörgö and Horváth (1988) for kernel density estimators, a Kiefer–Wolfowitz type result no longer suffices. In that case, results on the L_p-distance, between $\widehat{F}_n$ and F_n are more appropriate, such as the ones in Durot and Tocquet (2003) and Kulikov and Lopuhaä (2008).

In this paper, we extend the results in Durot and Tocquet (2003) and Kulikov and Lopuhaä (2008) to the general setting of Durot (2007). Our main result is a central limit theorem for the L_p-distance between $\widehat{\Lambda}_n$ and Λ_n, where Λ_n is a naive estimator for the primitive Λ of a monotone curve λ and $\widehat{\Lambda}_n$ is the LCM of Λ_n. As special cases we recover Theorem 5.2 in Durot and Tocquet (2003) and Theorem 2.1 in Kulikov and Lopuhaä (2008). Our approach requires another preliminary result, which might be of interest in itself, i.e., a limit process for a suitably scaled difference between $\widehat{\Lambda}_n$ and Λ_n. As special cases we recover Theorem 1 in Wang (1994), Theorem 4.1 in Durot and Tocquet (2003), and Theorem 1.1 in Kulikov and Lopuhaä (2006).

2. Main results

We consider the general setting in Durot (2007). Let $\lambda : [0, 1] \to \mathbb{R}$ be nonincreasing and assume that we have at hand a cadlag step estimator Λ_n of

$$\Lambda(t) = \int_0^t \lambda(u)\,\mathrm{d}u, \quad t \in [0, 1].$$

In the sequel we will make use of the following assumptions.

(A1) λ is strictly decreasing and twice continuously differentiable on $[0, 1]$ with $\inf_t |\lambda'(t)| > 0$.

(A2) Let B_n be either a Brownian motion or a Brownian bridge. There exists $q > 6$, $C_q > 0$, $L : [0, 1] \to \mathbb{R}$, and versions of $M_n = \Lambda_n - \Lambda$ and B_n such that

$$\mathbb{P}\left(n^{1-1/q} \sup_{t\in[0,1]} \left|M_n(t) - n^{-1/2} B_n \circ L(t)\right| > x\right) \le C_q x^{-q}$$

for all $x \in (0, n]$. Moreover, L is increasing and twice differentiable on $[0, 1]$, with $\sup_t |L''(t)| < \infty$ and $\inf_t |L'(t)| > 0$.

Note that this setup includes several statistical models, such as monotone density, monotone regression, and the monotone hazard model under random censoring, see Durot, (2007) [Section 3].

We consider the distance between Λ_n and its least concave majorant $\widehat{\Lambda}_n = \mathrm{CM}_{[0,1]}\Lambda_n$, where CM_I maps a function $h : \mathbb{R} \to \mathbb{R}$ into the least concave majorant of h on the interval $I \subset \mathbb{R}$. Consider the process

$$A_n(t) = n^{2/3}\left(\widehat{\Lambda}_n(t) - \Lambda_n(t)\right), \qquad t \in [0, 1], \tag{1}$$

and define

$$Z(t) = W(t) - t^2, \qquad \zeta(t) = [\mathrm{CM}_{\mathbb{R}} Z](t) - Z(t), \tag{2}$$

where W denotes a standard two-sided Brownian motion originating from zero. For each $t \in (0, 1)$ fixed and $t + c_2(t)sn^{-1/3} \in (0, 1)$, define

$$\zeta_{nt}(s) = c_1(t) A_n\left(t + c_2(t)sn^{-1/3}\right), \tag{3}$$

where

$$c_1(t) = \left(\frac{|\lambda'(t)|}{2L'(t)^2}\right)^{1/3}, \qquad c_2(t) = \left(\frac{4L'(t)}{|\lambda'(t)|^2}\right)^{1/3}. \tag{4}$$

Our first result is the following theorem, which extends Theorem 1.1 in Kulikov and Lopuhaä (2006).

Theorem 1. *Suppose that assumptions (A1)–(A2) are satisfied. Let ζ_{nt} and ζ be defined in (3) and (2). Then the process $\{\zeta_{nt}(s) : s \in \mathbb{R}\}$ converges in distribution to the process $\{\zeta(s) : s \in \mathbb{R}\}$ in $D(\mathbb{R})$, the space of cadlag function on $\mathbb{R}$.*

Note that as a particular case $\zeta_{nt}(0)$ converges weakly to $\zeta(0)$. In this way, we recover Theorem 1 in Wang (1994) and Theorem 4.1 in Durot and Tocquet (2003). The proof of Theorem 1 follows the line of reasoning in Kulikov and Lopuhaä (2006).

Let us briefly sketch the argument to prove Theorem 1. Note that $A_n = D_{[0,1]}[n^{2/3}\Lambda_n]$ and $\zeta = D_{\mathbb{R}}[Z]$, where $D_I h = \mathrm{CM}_I h - h$, for $h : \mathbb{R} \to \mathbb{R}$. Since D_I is a continuous mapping, the main idea is to apply the continuous mapping theorem to